Journal of Hydrology 519 (2014) 1052–1067

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain

Francesco Silvestro *, Nicola Rebora

CIMA Research Foundation, Savona, Italy[1]

ARTICLE INFO

Article history:
Received 12 February 2014
Received in revised form 18 July 2014
Accepted 20 July 2014
Available online 30 July 2014
This manuscript was handled by Geoff Syme, Editor-in-chief

Keywords:
Flood
Forecast
Probabilistic
Ensemble
Expert precipitation forecast
Small basins

SUMMARY

One of the main difficulties that flood forecasters are faced with is evaluating how errors and uncertainties in forecasted precipitation propagate into streamflow forecast. These errors, must be combined with the effects of different initial soil moisture conditions that generally have a significant impact on the final results of a flood forecast. This is further complicated by the fact that a probabilistic approach is needed, especially when small and medium size basins are considered (the variability of the streamflow scenarios is in fact strongly influenced by the aforementioned factors). Moreover, the ensemble size is a degree of freedom when a precipitation downscaling algorithm is part of the forecast chain. In fact, a change of ensemble size could lead to different final results once the other inputs and parameters are fixed. In this work, a series of synthetic experiments have been designed and implemented to test an operational probabilistic flood forecast system in order to augment the knowledge of how streamflow forecasts can be affected by errors and uncertainties associated with the three aforementioned elements: forecasted rainfall, soil moisture initial conditions, and ensemble size.

1. Introduction

A reliable operational flood forecasting system is important if streamflow, possible flooding and its effects are to be predicted with sufficient lead time to allow for appropriate actions. This is a particularly challenging task in the case of small to medium sized basins with drainage areas of order of magnitude from 10 to 10,000 km^2 or less (Siccardi et al., 2005). In this context, two main issues arise: (i) the need to base the flood prediction system on rainfall forecast (because of the fast response time of these catchments), and (ii) the need to follow a probabilistic approach (because of the inconsistencies between meteorological modeling and hydrologic response, Ferraris et al., 2002).

Such forecasting systems, as a basic option, use an atmospheric model and a hydrological prediction system. When dealing with small catchments however, a downscaling module is necessary to account for uncertainties and inconsistencies in the rainfall predicted by meteorological models (Mascaro et al., 2010; Siccardi et al., 2005).

Each component of the system is affected by uncertainty and these uncertainties propagate and amplify from the beginning of the chain (the atmospheric models) to the discharge or its transformation into river level and eventually the corresponding flood scenarios.

Many works have been devoted to dealing with errors and uncertainties related to: (i) precipitation estimation (e.g. Germann et al., 2009), (ii) hydrological model parameterization (e.g. Carpenter and Georgakakos, 2006; Vrugt et al., 2006), (iii) soil moisture initial conditions (e.g. Brocca et al., 2011; Tramblay et al., 2012), and (iv) forecasted rainfall (e.g. Siccardi et al., 2005). Zappa et al. (2011) analyzed the superposition of different sources of uncertainties. Mascaro et al. (2010) analyzed how ensemble precipitation errors affect streamflow simulations using a multifractal rainfall downscaling model coupled with a distributed hydrological model. However, most of these works analyzed a particular source of uncertainty, with no reference to the multi-catchment approach or to the influence of sample size on the results of probabilistic chains. In addition, the errors related to the quantitative use of expert precipitation forecasts were never discussed.

In this work, a probabilistic chain used operationally for flood forecasts on small and medium-sized basins in Liguria, Italy is considered (Silvestro et al., 2011, 2012) and several synthetic experimental suites have been designed and implemented in order to provide answers to practical and operational questions. That is,

* Corresponding author. Address: CIMA Research Foundation, University Campus, Armando Magliotto, 2. 17100 Savona, Italy. Tel.: +39 019230271; fax. +39 01923027240.
E-mail address: francesco.silvestro@cimafoundation.org (F. Silvestro).
[1] www.cimafoundation.org

http://dx.doi.org/10.1016/j.jhydrol.2014.07.042
0022-1694/

how errors in the forecasted rainfall propagate and affects the final results of the flood forecasting chain and how these effects change and combine with the soil moisture initial conditions. Further, since we considered a probabilistic system, an analysis of the effects related to the ensemble size is also done.

For simplicity, the uncertainties related to the hydrological model parameters (Beven and Binley, 1992; Liu et al., 2005; Zappa et al., 2011) have been neglected. These parameters can have considerable effects on flood simulations, but they are generally negligible with respect to those related to the rainfall forecasts and initial wetting conditions (Mascaro et al., 2010; Zappa et al., 2011) especially when dealing with small basins. In small basins, errors on precipitation localization or bad estimation of rainfall amounts at small spatial–temporal scales can lead to huge errors in flood forecasts. Moreover, in the case presented in this paper, the considered hydrological model is quite consolidated in the test region and the suitability of the operational parameters set is constantly verified comparing simulations with the streamflow observations.

There are three main peculiarities of the flood forecast chain under consideration: (i) the forecasted precipitation is quantitatively predicted by an expert forecaster and not by a Numerical Weather Prediction System (NPWS), (ii) the expert forecast is downscaled in order to properly account for the uncertainties on small catchments, and (iii) the forecast is carried out following two different approaches depending on the dimension of the basins: single-site and multi-catchment (Siccardi et al., 2005).

The forecasted rainfall is stochastically downscaled with an algorithm based on Rebora et al. (2006) in order to generate an ensemble of rainfall fields of fine spatial–temporal scales [1 km - 0.5 h]. As a consequence, we seek to investigate the following issues: (a) how errors in the precipitation forecast affect the rainfall ensemble (b) how these errors propagate into the streamflow ensemble and finally, (c) how these errors are amplified or reduced by the different soil moisture initial conditions.

The manuscript is organized as follows: Section 2 describes the context and the flood forecasting chain, Section 3 shows the design of the experiments while in Section 4 the results are reported, and finally the paper concludes in Section 5 with the discussion and conclusions.

2. Case study

2.1. The expert quantitative precipitation forecast (EQPF)

As described in Silvestro et al. (2011) the precipitation forecast for the Liguria Region is provided by a number of Numerical Weather Prediction Models (NWPM) and interpreted by expert meteorologists (see Table 1). The Liguria Region, Italy is divided into five alert sub-regions that are considered homogenous from the hydrological response point of view, and also from their meteo-climatic characteristics (Fig. 1). Each alert sub-region has an area of the order of 10^3 km^2 and contains a number (around 20) of modeled small sized basins (O(Area) 10-10^2 km^2) on a complex topography with steep valleys and concentration times ranging from 1 to 6 h. Few basins in this Region have larger areas and concentration times (O(Area) 10^2-10^3 km^2). A high percentage of the territory is covered by forests or by lawns with shrubs (about 70%) and the main urban areas and towns are established along the coast, often at the mouth of creeks and rivers.

Table 1
NWPM used by the meteorologists of HMFC to carry out the expert quantitative precipitation forecast.

Model Name	Spatial resolution (km)	Type
ECMWF	30	Global scale
COSMO-LAMI	7	Mesoscale
BOLAM	10	Mesoscale
MOLOCH	2	Regional scale

The experts of the Hydro-Meteorological Functional Centre of Liguria Region (HMFC) merge the output of the different meteorological models (the so called "poor man ensemble" e.g. Arribas et al., 2005) with their own experience and provide quantitative precipitation forecasts for the alert sub-regions over predefined time windows. For each alert sub-region, a different quantitative precipitation forecast is made.

We briefly report how the rainfall forecasts are specifically made and used operationally for feeding the flood forecast chain (for further details see Silvestro et al., 2011).

Basically, three quantities are provided for a given rainfall event:

1. The maximum average precipitation in a time window of 12 h for each homogeneous sub-region (named P_{12}). In order to define this quantity, a certain number of meteorological models (Table 1) are analyzed in order to identify the 12 h time window when the maximum precipitation amount is expected. The starting time of the time window is identified as t_{12}. The 12 h window is chosen since it is the typical length of an extreme rainfall event in the area.
2. Once the P_{12} is defined, the meteorologist estimates the rainfall amount (Pb) that is expected between the reference starting time of the forecast (t_0) and the start of the 12 h window (t_{12}) associated with the maximum volume. The time window where Pb is estimated is then $\tau_b = t_{12} - t_0$. This is done for each sub-region. The reference start time (t_0) is the time at which the meteorologist starts his forecast and this is usually 00:00 h on the day he makes the prediction.
3. The third parameter is the maximum precipitation amount forecasted in a time window of 3 h and on areas of approximately 10^2 km^2 (named P_3), that is, on boxes of 10×10 km^2. This number gives an idea of the local intensity of the forecasted event; high values mean possibly critical situations for basins with areas in the range of 10^1–10^2 km^2. It also indicates how much the precipitation volume, defined by P_{12}, tends to be concentrated in localized areas. Generally, a P_3 value is given for each sub-region, but it can also be estimated as a single P_3 value for the entire regional territory. Meteorologists are confident to issue quantitative precipitation forecast up to spatial–temporal scale of 10^2 km^2 – 3 h, which is coherent with the scale of dimension and time of concentration of most of the modeled basins.

Given that P_{12} and P_3 are estimated on different reference areas, the value of P_3 can be larger than P_{12}. In fact, the condition that must be verified is on the total volumes:

$$P_3 \times A_3 \leq P_{12} \times A_{12} \quad (1)$$

where A_3 is the 10^2 km^2 box and A_{12} the area of the homogeneous region.

An automatic software calculates the P_3 and P_{12} values for each model of Table 1 and merges them but the forecasters can decide to modify these values based on their confidence with the used NWPM and their subjective assessment of the occurring event.

2.2. The flood forecasting chain operational at the HMFC (FFC)

The EQPF is used as the input for the operational probabilistic hydro-meteorological forecasting chain (Siccardi et al., 2005;

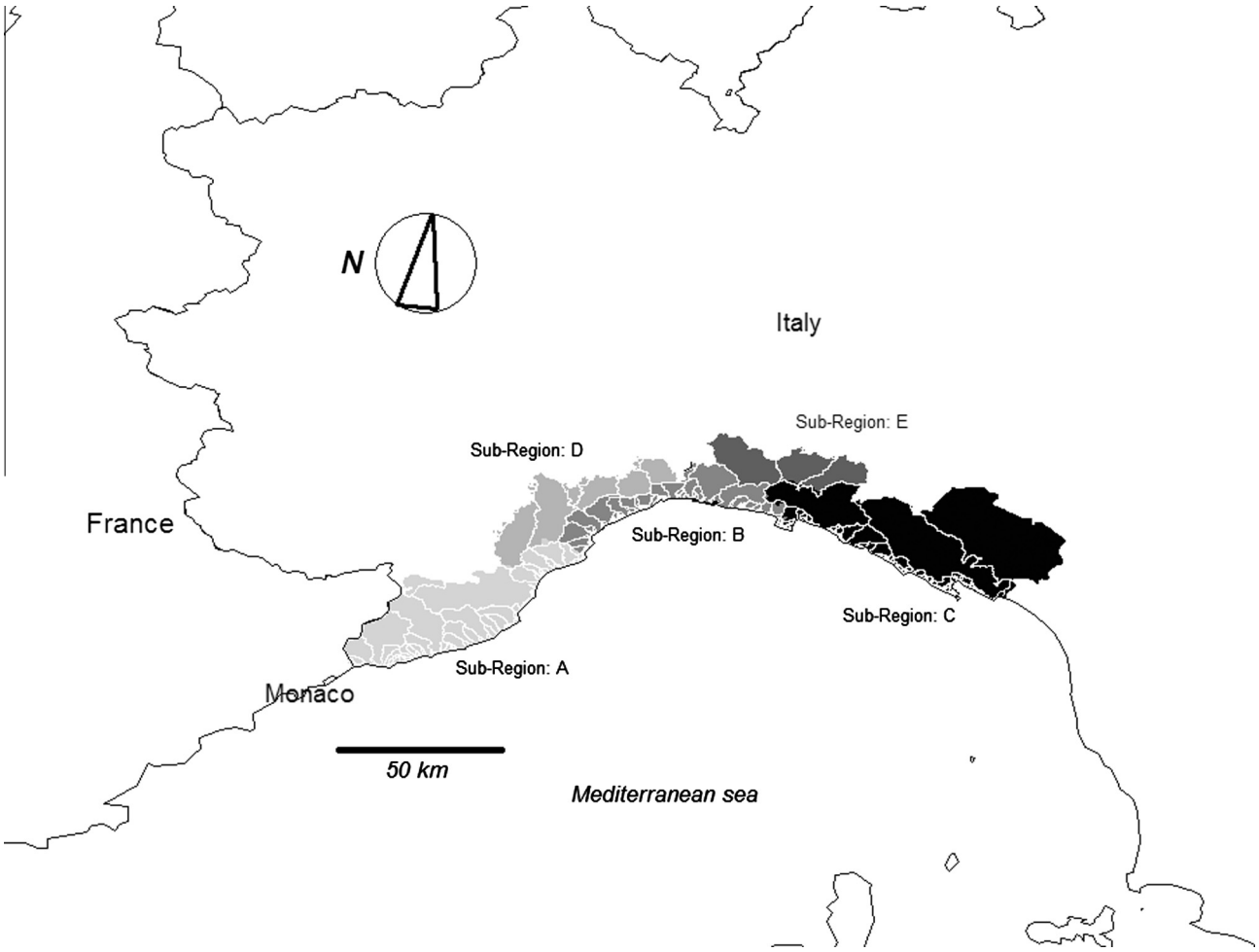

Fig. 1. The Liguria Region, Italy is divided into five alert sub-regions as shown. The main basins are shown with white outlines.

Silvestro et al., 2011). The other components of the hydro-meteorological chain (Fig. 2) are the downscaling module RainFARM (Rebora et al., 2006) adapted for such an input, and the hydrological model, DRiFt (Giannoni et al., 2000, 2003; Gabellani et al., 2008).

DRiFt is an event scale semi-distributed model; the model is focused on the efficient description of the drainage system in its essential parts: hillslopes and channel networks. These are addressed by using two kinematic scales which determine the basis for the geomorphologic response of the basin. The propagation of water driven by gravity in the top soil layer is described and in this way an auto initialization of the model is reproduced between events. The schematization is valid and applicable when the simulation period is not too long and the evapotranspiration does not become crucial in the mass balance equation.

The basin is discretized into cells based on a Digital Elevation Model and two kinematic parameters, the channel flow velocity and the hillslope flow velocity, are set for cells classified as channel and hillslope respectively; these two parameters together with the flow paths define the routing time for each cell. The runoff estimated at the cell scale is routed to the outlet section without accounting for the storage in the channels and re-infiltration. The implemented infiltration scheme (Gabellani et al., 2008) models the soil moisture as the ratio between the actual and the maximum soil storage capacity of a given cell. The soil moisture initial conditions (SM) are estimated using an Antecedent Precipitation Index (API) methodology based on the precipitation that occurred in the month that precedes the event. The API is related to the initial soil moisture that assumes values in the range (0, 1) in the model. The model is widely applied in the considered area, given its use for flood forecasting. The parameters are generally calibrated with the aim of reproducing the time-to-peak and the peak of streamflow values.

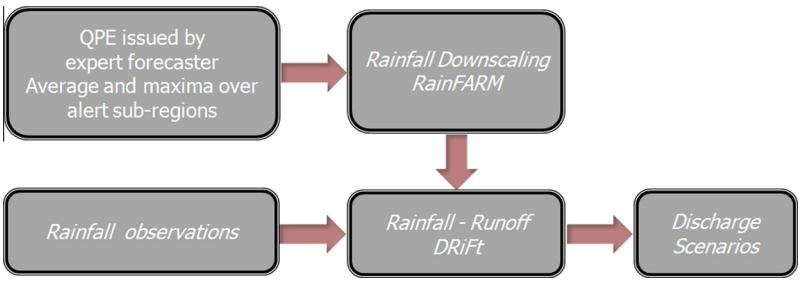

Fig. 2. Scheme of a generic probabilistic flood forecast chain based on a downscaling system and an hydrological model fed with observed and forecasted rainfall.

The downscaling module (RainFARM) produces an ensemble of high-resolution precipitation fields by preserving the information at the large scale derived from a quantitative precipitation prediction. In this particular configuration, it is designed to preserve the values of P_{12} and P_3 (see previous sections) on each sub-region. The model parameters (α: spatial spectral slope; β: temporal spectral slope) are related to P_{12} and P_3 using the approach followed by Silvestro et al. (2011) which defined a look-up table that returns a parameter set for each couple of possible values of P_{12} and P_3,

As a consequence, RainFARM is capable of generating an ensemble of size N_e (with N_e of order of magnitude 100–500 members) of small scale (e.g. spatial resolution: 1 km, temporal resolution: 0.5 h) precipitation fields of 12 h duration that are consistent with radar observations of mid-latitude precipitation events (Rebora et al., 2006). Generally, as the ratio between P_3 and P_{12} increases, RainFARM tends to concentrate the precipitation on small spatial–temporal scales accentuating the convective characteristic of the event. Each of the N_e downscaled precipitation fields is then used as the input for the hydrological model, down to scales of 1 km^2, in order to generate an ensemble (with size N_e) of streamflow scenarios on each catchment. The results are post-processed producing flood predictions following two different approaches: the single-site and the multi-catchment, as described by Siccardi et al. (2005).

The two approaches are differentiated based on the definition of two pairs of quantities: (i) the meteorological reliable spatial–temporal scales (l_{met}, t_{met}), which are the scales at which a reliable quantitative precipitation forecast is expected, and (ii) the scales of the hydrological processes (l_{hydro}, t_{hydro}) that depend on the dimension and characteristics of the affected basins. Operationally, in the case of the Liguria Region's forecasting chain, the domain of l_{met} is assumed to be the size of the alert sub-region that has an area of the order of 10^3 km^2 while $t_{\mathrm{met}} = 12$ h. Following Siccardi et al. (2005), the single site approach can be adopted when $O(l_{\mathrm{met}}/l_{\mathrm{hydro}}) < 10^2$ while, when $O(l_{\mathrm{met}}/l_{\mathrm{hydro}}) \geqslant 10^2$ the multi-catchment approach must be adopted.

In the studied FFC, the single site approach is operationally applied to all the modeled basins, but is mainly useful for those basins that have an area larger than 200–300 km^2 since $l_{\mathrm{met}} \sim 10^3$ (Silvestro et al., 2011). In this case, the probability that a certain flow threshold (or the flow with a given return period, T) could be exceeded is directly evaluated for the single basin. For smaller basins the multi-catchment approach is applied. The forecasting procedure does not allow for any discrimination between different spatial localizations within the domain of l_{met}. We thus cannot consider every single basin as an independent entity, but we consider together all the basins belonging to the same area represented by l_{met}.

For each alert sub-region a different forecast is produced in terms of a probability curve. The procedure evaluates the probability that, in at least one basin belonging to that specific alert sub-region, the flow with a given return period T will be exceeded. The procedure does not specify which basin will experience the particular flow, due to the uncertainty associated with the meteorological forecast. This procedure represents a possible solution in the case of flood forecasts for small ($O(A) < 100$ km^2) basins.

3. Design of experiments

3.1. Experiment 1: Combined effects of errors in P_{12} and SM

The first experiment has three objectives: (i) to investigate the effects of the SM conditions on the streamflow forecast ensemble for events of varying severity; (ii) to investigate how errors in the precipitation forecast (i.e. P_{12}) propagates into the streamflow forecast ensemble; (iii) the combined effects of points i and ii.

The distribution of the error on P_{12} is unknown so we considered it uniformly distributed similarly to what was done by Carpenter and Georgakakos (2006) for modeling parameters' uncertainty. Error on P_{12} is due to several reasons, but mainly related to the errors on the Quantitative Precipitation Forecast (QPF) of NWPS and to the errors made by the meteorologists in interpreting the different NWPS.

In this first experiment, P_3 is not considered as a variable and a set of typical values of the two RainFARM parameters is assumed ($\alpha = 1.7$; $\beta = 1.2$).

The steps of the experiment are as follows:

1. The values of the RainFARM parameters are fixed using typical values (Rebora et al., 2006).
2. A series of realistic values of P_{12} have been considered (from 20 to 150 mm). The range was defined based on the experience of the forecasters and on P_{12} values which were observed during the most severe past events.
3. Three typical SM have been considered (dry SM = 0.2; normal SM = 0.5; wet SM = 0.7). As shown in Section 2, DRiFt representation of soil humidity ranges from 0 to 1 and from the cold to the hot season almost the entire range is experienced; we thus used the aforementioned values as representative of the possible soil moisture initial conditions (dry, normal, wet soil).
4. The reference flood forecasts (benchmark, also called perfect forecast) were built by using for each P_{12} value and each SM value a relatively large ensemble size ($N_e = 400$), so that N_e has negligible influence on the results. In practice, the perfect forecast is done assuming that the rainfall input P_{12} is the true value that is forecasted without errors.
5. N_{er} percentage errors (positive and negative) have been associated with the P_{12} values aiming to reproduce possible estimation errors in an operational context, and the flood forecasts were then generated. Errors from 5% to 30% have been considered. For example, Err + 5% means that a value of $P_{12} + 0.05 * P_{12}$ is used as input for the forecast.

Finally, for each pair of P_{12} and SM, $N_{er} + 1$ flood forecasts made by N_e streamflow scenarios for each modeled section are available for analysis and post-processing.

In practice P_{12} and SM have been varied in order to simulate a set of possible precipitation forecasts and soil moisture initial conditions, then a set of percentage errors have been applied to P_{12} in order to simulate the uncertainty of the precipitation forecast.

3.2. Experiment 2: Combined effects of errors in P_{12}, P_3, SM

The second experiment has the objective of evaluating the relative effect of P_{12} and P_3 for the different SM under consideration.

The steps of the experiment are as follows:

1. Three values of P_{12} were considered based on experience, as representative of small, medium and large volumes of rainfall typical of the considered environment ($P_{12} = 20$, 50, 100 mm).
2. A set of realistic P_3 values based on historical events were identified (from 10 to 100 mm).
3. Three typical SM were considered (dry SM = 0.2; normal SM = 0.5; wet SM = 0.7).
4. The values of P_{12}, P_3 and SM were combined in order to run the FFC and produce the flood forecasts with $N_e = 400$.

This experiment simulates the effects of possible errors and inaccuracies made by the forecaster in predicting P_3 values. The errors in predicting P_3 are potentially large because the temporal and spatial scales are finer with respect to P_{12}. For this reason, we combined each P_{12} value with all the considered P_3 values

and without assuming a percentage error. In this case all the three quantities P_{12}, P_3 and SM have been varied producing a set of possible forecasts that differ in terms of volume of precipitation and type of rainfall event, the ratio P_3/P_{12} determines if the rainfall tends to be uniformly distributed in space and time or if it tends to be concentrated on small spatial–temporal scales.

For each combination of P_{12}, P_3 and SM a flood forecast made by N_e streamflow scenarios for each modeled section is available for analysis and post-processing.

3.3. Experiment 3: Uncertainty related to the ensemble size

The third experiment has the objective of evaluating the impact of the ensemble size, N_e on the final results of the forecast. It is in fact, evidently impossible to represent all the possible occurrences with RainFARM as with a generic downscaling system in a real-time application. In experiment 1 and 2 we assumed a very high N_e value in order to avoid its impact on the results, in experiment 3 we try to answer to the following question: what is a reasonable value of N_e such that repeating the same forecast two or more times the results have negligible differences?

The answer is that, generally, this number varies for different values of precipitation input and for different soil moisture initial conditions. In this experiment, we consider critical conditions for carrying out the analysis. The results of Experiment 2, as will be shown in Section 4.2, provide evidence that very high values of P_{12} would not be the right choice since it is unlikely to have correspondingly high values of the ratio P_3/P_{12}; on the other hand, too small values would not result in particularly severe events and they are not physically compatible with very high values of P_3. Based on these considerations values for P_{12} and P_3 were chosen in order to have a significant total volume of precipitation and a high variability of the peak flows.

After fixing the values of P_{12} and P_3, the ensemble size N_e was varied (N_e = 5, 10, 15,...,250), and for each value of N_e a number M of forecasts (realizations of the forecast) were generated, with M being considerably large (M = 100). Once we fixed the values of P_{12}, P_3, SM and ensemble size N_e we thus have M possible realizations of the flood forecast which should be statistically equivalent, in operational applications this condition could be not satisfied since N_e is finite.

Considering the peak flows of a generic basin, it is possible to calculate the following quantities:

1. The mean peak flow for each of the M ensembles of size N_e (QM_e).
2. The standard deviation of peak flows for each of the M ensembles of size N_e (σQ_e).
3. The standard deviation of the QM_e for each ensemble size N_e (σQM_e); this quantity measures how the mean peak flow varies among the M ensembles of the same size.
4. The sample standard deviation of the standard deviation of peak flows $Q_{p,e}$ for each ensemble size N_e ($\sigma(\sigma Q_e)$); this quantity measures how the standard deviation of peak flows varies among the M ensembles of the same size.

These statistics are then used to build graphs in order to answer to the investigations of the experiment.

4. Results

4.1. Experiment 1

The outputs of the FFC were post-processed in order to produce graphs that synthesize the results. Since we deal with a probabilistic FFC box plots are particularly useful to present the outcomes.

In Fig. 3, the results obtained using the perfect forecast (where N_e = 400; subscript 0) are shown for basins of different areas and for different values of P_{12} applied to all the alert sub-regions. We indicate with the perfect forecast the resulting flood forecasts using P_{12} values without the application of any percentage error. The considered basins belong to sub-regions B and C. The wet SM (SM = 0.7) is considered in order to reduce the non-linearity of the runoff process and as consequence the impact on the variability of the streamflow scenarios. The transformation from rainfall to runoff through a generic infiltration algorithm is in fact non-linear and characterized by a threshold effect: the runoff can be null for a number of simulation time steps even if rainfall is larger than 0, becoming suddenly positive when the soil is sufficiently saturated. This is a good representation of the actual physical behavior of the soil. The box-plots are built after the normalization of each peak flow (Qp_0) with the average of the peak flows (Qp_{0M}).

The graphs in Fig. 3 show that the variability of the flow for the first three basins (area 10–300 km^2) is quite similar, while the integration effect is a little more evident for the basin with area of order 10^3 km^2. As shown in Silvestro et al. (2011), the minimum area of the basin to apply the single-site approach in the adopted system is about 200 km^2, therefore hereafter the basin corresponding to the low left-hand corner of Fig. 3 is considered to show the results in the case of the single-site approach. As shown, the median is generally different from the mean; this indicates that the peak flows are not normally distributed (Ferraris et al., 2002; Alfieri et al., 2012). The upper whiskers are larger than the lower whiskers in part because of the skewness of the distribution and also because of the normalization effect.

The effect on larger basins is a general reduction of the variability of the forecast. Regarding the multi-catchment approach, no particular reasons exist for choosing a certain sub-region and not another for the purpose of this analysis, in fact the forecast is done with the same method and same values for all the sub-regions. In all cases, basins ranging in area from 20 to at least 150 km^2 are present with morphologic and infiltration characteristics that do not change dramatically from one site to another. Sub-region C is only considered because it includes the test basin used for the single site approach (see Fig. 1).

4.1.1. Single site

The box plots are shown for the three considered SM initial conditions (Fig. 4) for the selected basin. For low values of P_{12} the variability is high (the whiskers assume values 6–8 times the Qp_{0M}), but in general the results are not considered to be significant because the flow values are low. With the increase of P_{12}, the box (interquartile) increases amplitude, this is due to the non-linearity of the runoff process. For very high P_{12}, both boxes and whiskers decrease, in this case the basin response tends to be linear and the variability is due to rainfall scenarios only.

This behavior is more relevant for wet initial conditions. In fact, for the same value of P_{12} both boxes and whiskers tend to be narrower when the SM increases.

In order to give an idea of how the results can vary as a function of basin area, we considered the four basins of Fig. 3 and three significant P_{12} values to build the plots in Fig. 5. The variability generally decreases with the increase of both drainage area and SM, but similar to what was deduced from Fig. 3, the decreasing variability with increased drainage area is evident in the larger basin (its area is of the same order of magnitude of the sub-region area).

The flood forecasts with the associated percentage errors were then considered. The mean peak flow of the forecast with a certain error applied to P_{12} (Qp_{err}) is compared with the one obtained by the perfect forecast; the results are presented in terms of percentage errors. The three subplots of Fig. 6 have the P_{12} on the x-axis and the relative error of the mean peak flow on the y-axis, each line

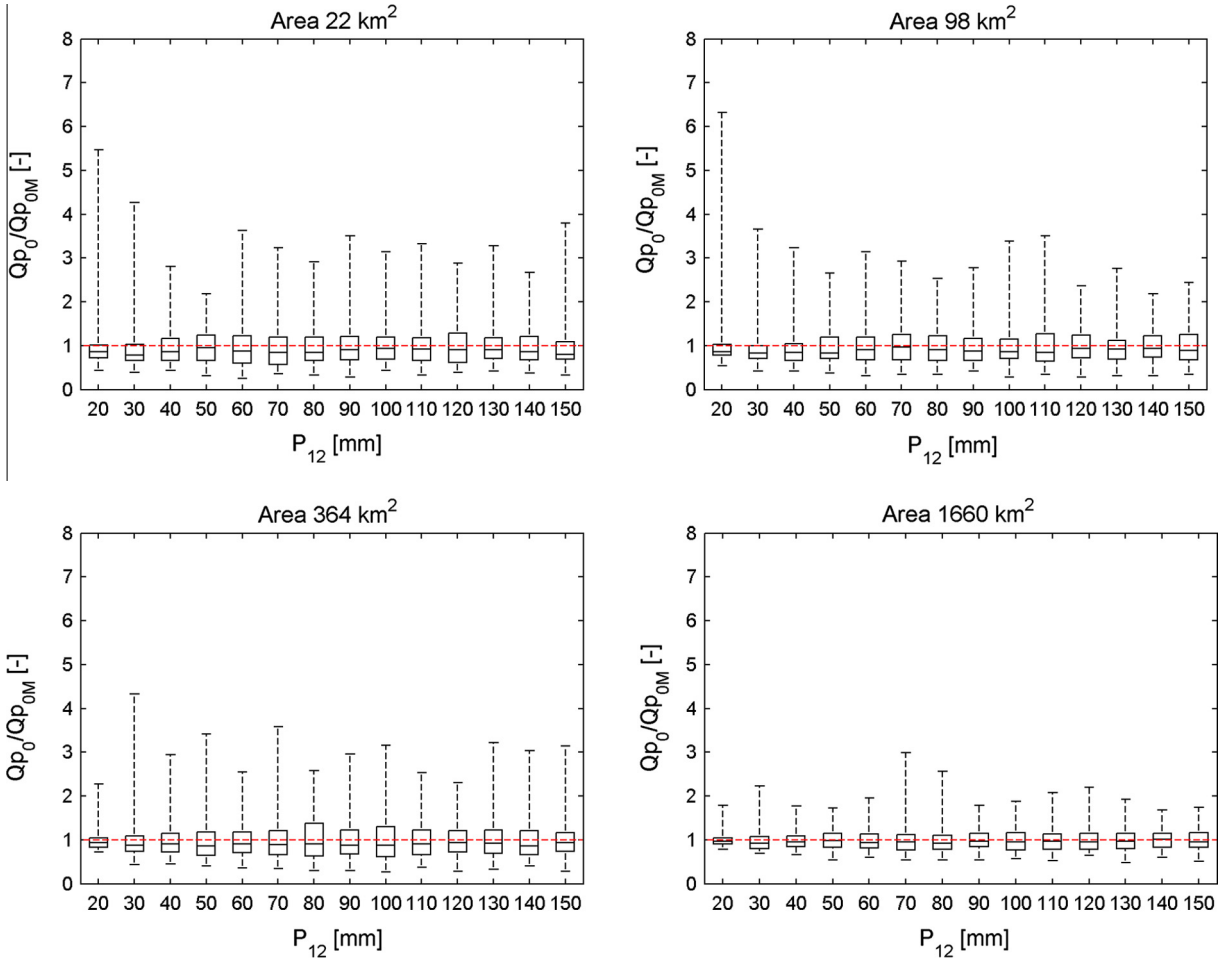

Fig. 3. Box plots of the peak flows (Qp_0) normalized with the average of the peak flows (Qp_{0M}) for different basins and different values of P_{12} (SM = 0.7).

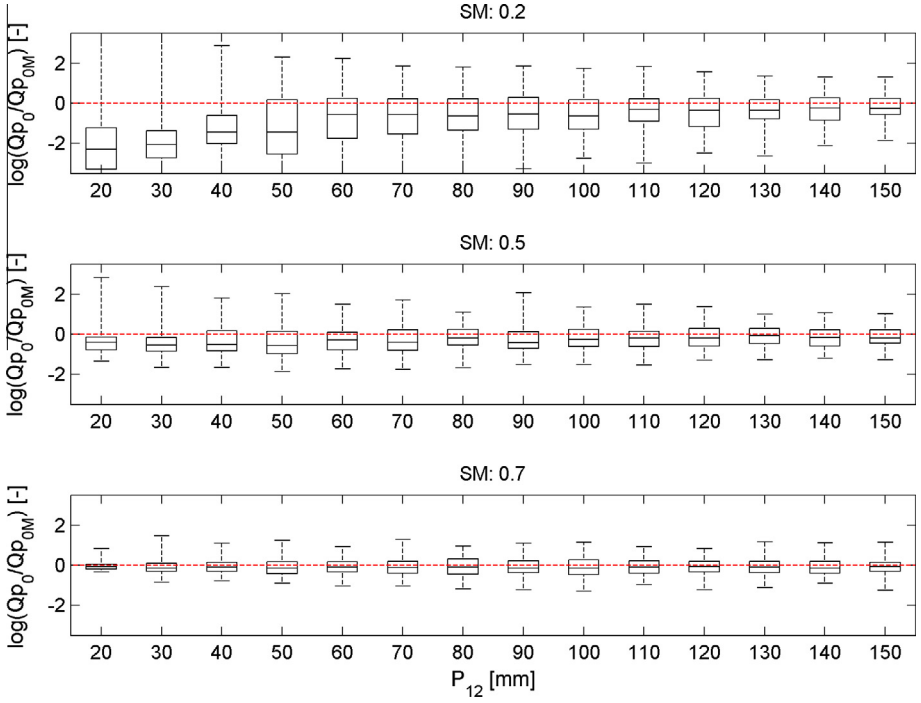

Fig. 4. Reference basin (364 km^2). Box plots of the peak flows (Qp_0) normalized with the average of the peak flows (Qp_{0M}) for different values of P_{12} and SM. Log scale is used on the *y*-axis.

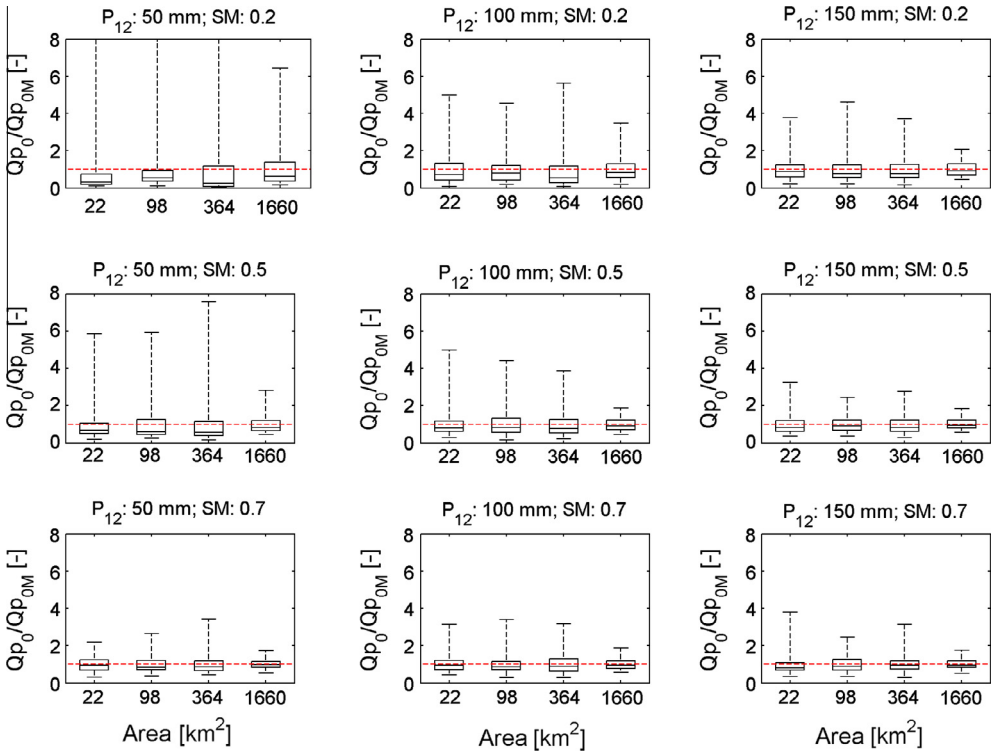

Fig. 5. Box plots of the peak flows (Qp_0) normalized with the average of the peak flows (Qp_0 M) for basins of different drainage area as a function of: (i) Pf_{12} values, and (ii) SM values.

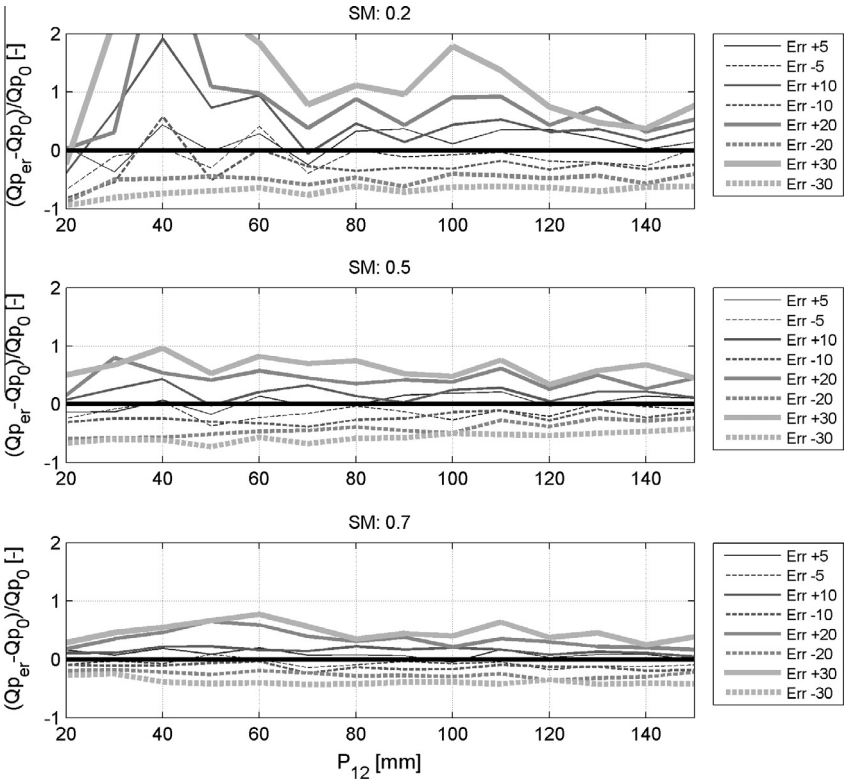

Fig. 6. Reference basin (364 km^2). Relative errors of the mean peak flow as a function of P_{12}, of the percentage error on Pf_{12} and of the soil moisture condition SM. The error is calculated with respect to the perfect forecast (zero % error on P_{12}).

corresponds to a percentage error of the P_{12}; each subplot corresponds to a different initial SM condition. The great non-linear effect of the runoff process is evident in the case of the initially dry soil. For SM = 0.2, in general, a certain error on P_{12} is amplified when the related streamflow prediction is considered. For example, for P_{12} = 80 mm, an error of 30% on precipitation leads to an

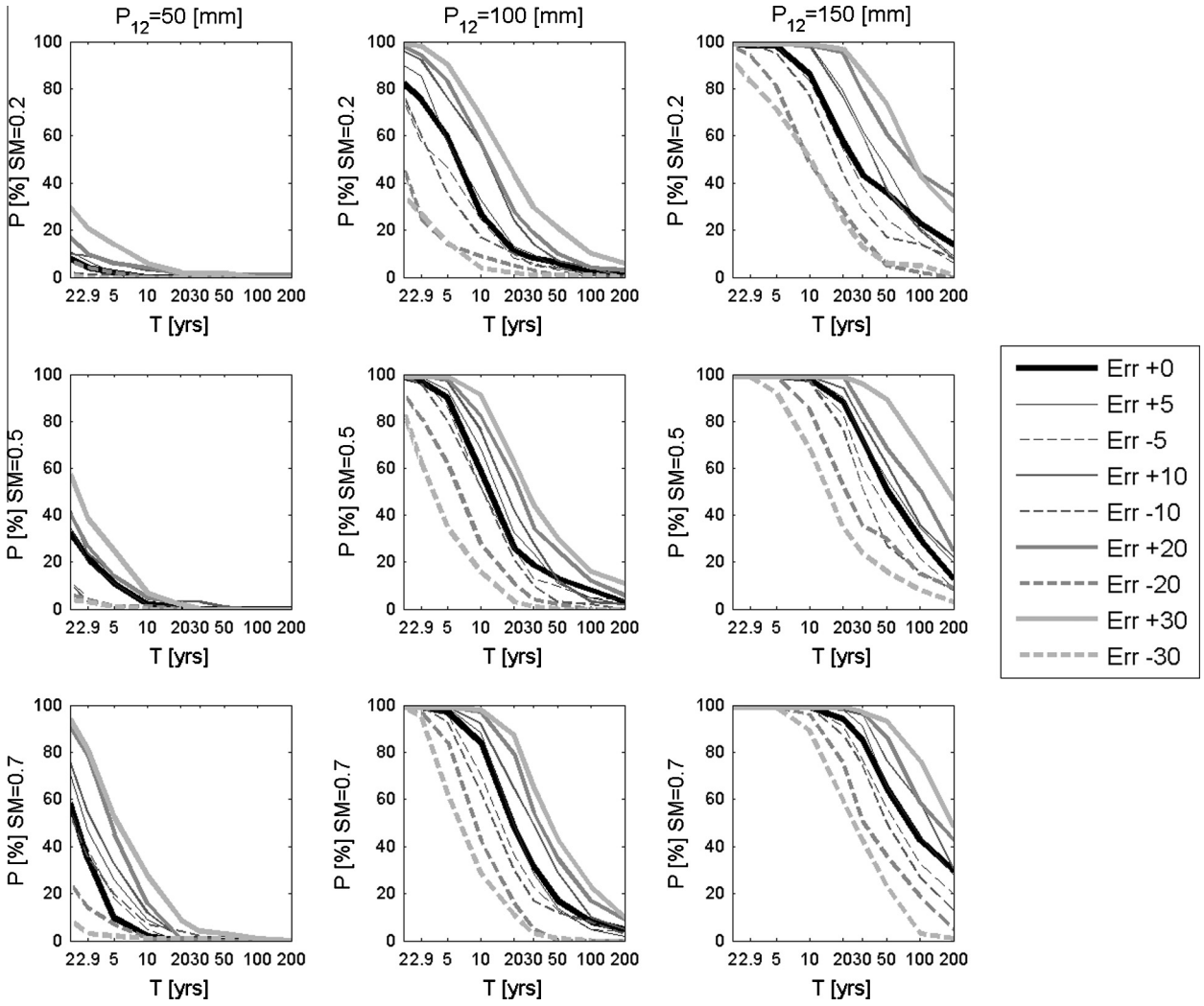

Fig. 7. Probability of exceedance curves for sub-region *C* as a function of the percentage error on P_{12} and of the soil moisture conditions, SM. P_{12} = 50, 100, 150 mm are considered.

Table 2
Peak flows for fixed return period of the basins belonging to sub-region *C*. Q are in m^3/s.

Basin code	Area (km^2)	Q 2	Q 2.9	Q 5	Q 10	Q 20	Q 30	Q 50	Q 100	Q 200
1	315	279	349	452	624	859	1013	1212	1484	1753
2	202	200	250	324	447	616	726	868	1063	1256
3	26	59	74	96	133	183	216	258	316	373
4	342	337	421	546	754	1037	1222	1463	1791	2116
5	43	98	123	159	220	302	356	426	522	617
6	297	461	577	747	1031	1418	1673	2002	2451	2895
7	25	58	73	94	130	179	211	253	310	366
8	165	209	261	338	467	643	758	906	1110	1311
9	153	162	202	262	361	497	586	702	859	1015
10	203	245	307	397	548	754	889	1064	1303	1539
11	549	481	601	778	1075	1479	1743	2086	2555	3018
12	56	116	145	188	260	358	430	510	630	740
13	25	64	80	104	144	198	233	279	342	404
14	129	254	317	411	567	781	995	1190	1458	1722
15	102	226	283	367	506	696	821	982	1203	1421
16	78	98	123	159	219	302	355	425	521	615
Test basin	364	543	679	879	1214	1670	1968	2356	2885	3408

error of approximately 100% on the peak flow when SM = 0.2. This effect decreases when SM increases, and for SM = 0.7 the basin behaves almost in a linear way: a certain error on precipitation propagates with a reduced amplification of streamflow values.

4.1.2. Multi-catchment

The effects of the uncertainty of P_{12} in the case of the multi-catchment approach were then considered. So the question is, how does the probability curve of a certain homogeneous region

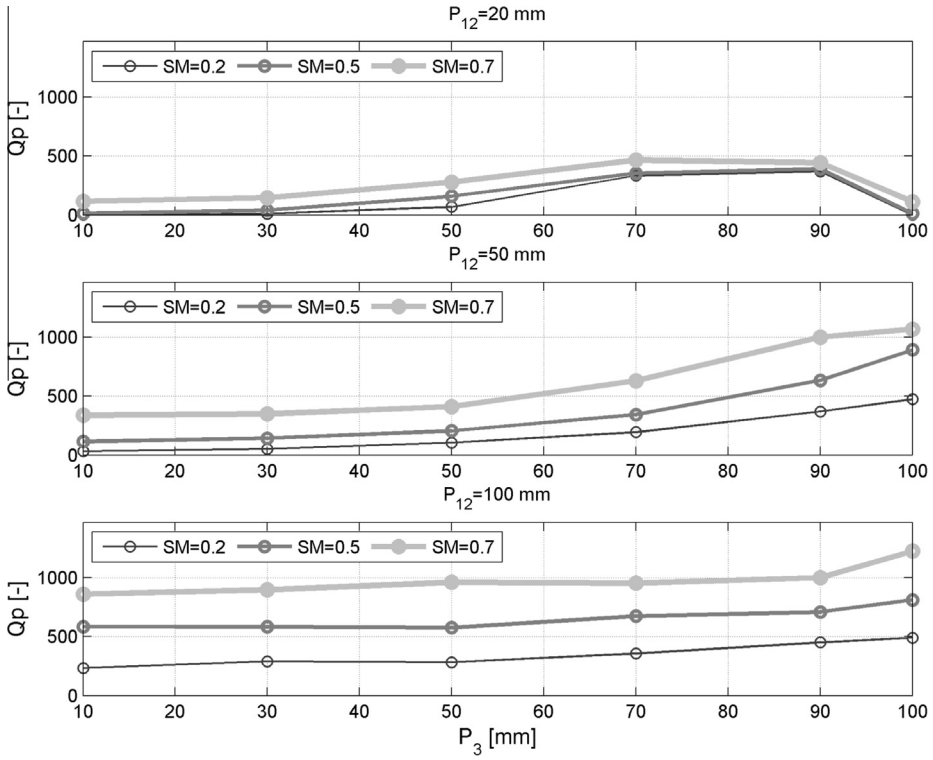

Fig. 8. Reference basin (364 km^2). Value of the mean peak flow for fixed values of P_{12} (20, 50, 100 mm) as a function of P_3 and of the soil moisture conditions, SM.

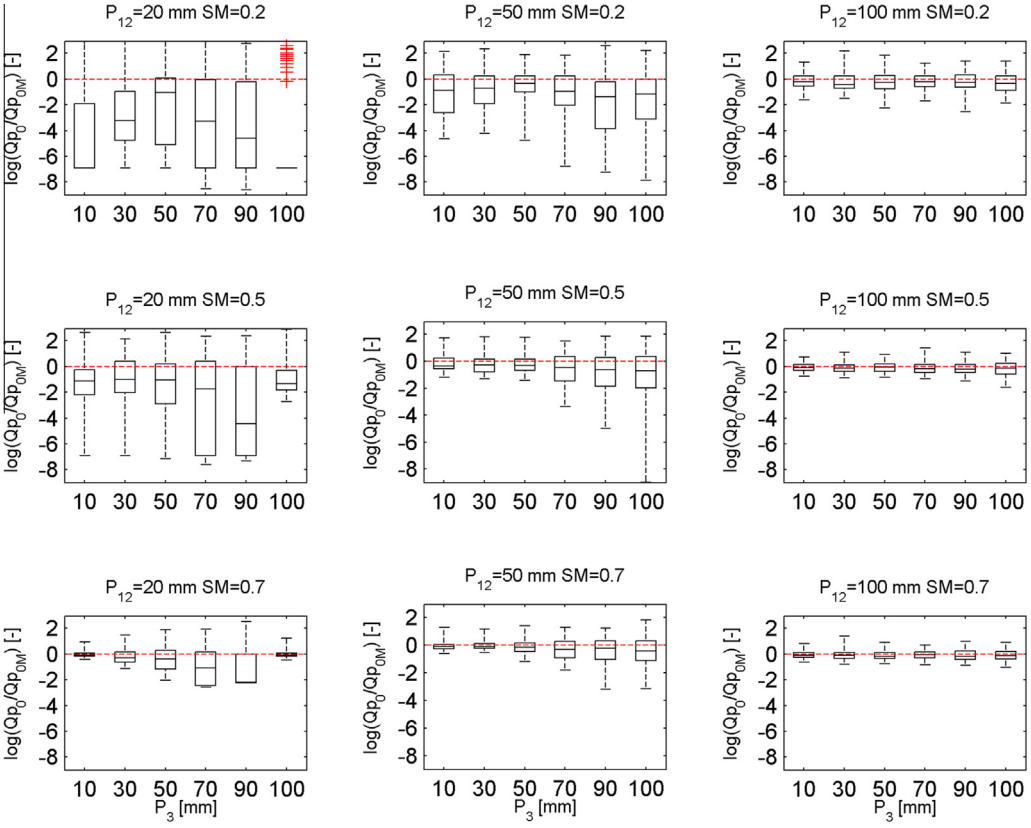

Fig. 9. Reference basin (364 km^2). P_{12} = 20, 50, 100 mm are considered. Box plots of the peak flows (Qp_0) normalized with the average of the peak flows (Qp_{0M}) for different values of P_3 and different values of SM. Log scale is used on y-axis.

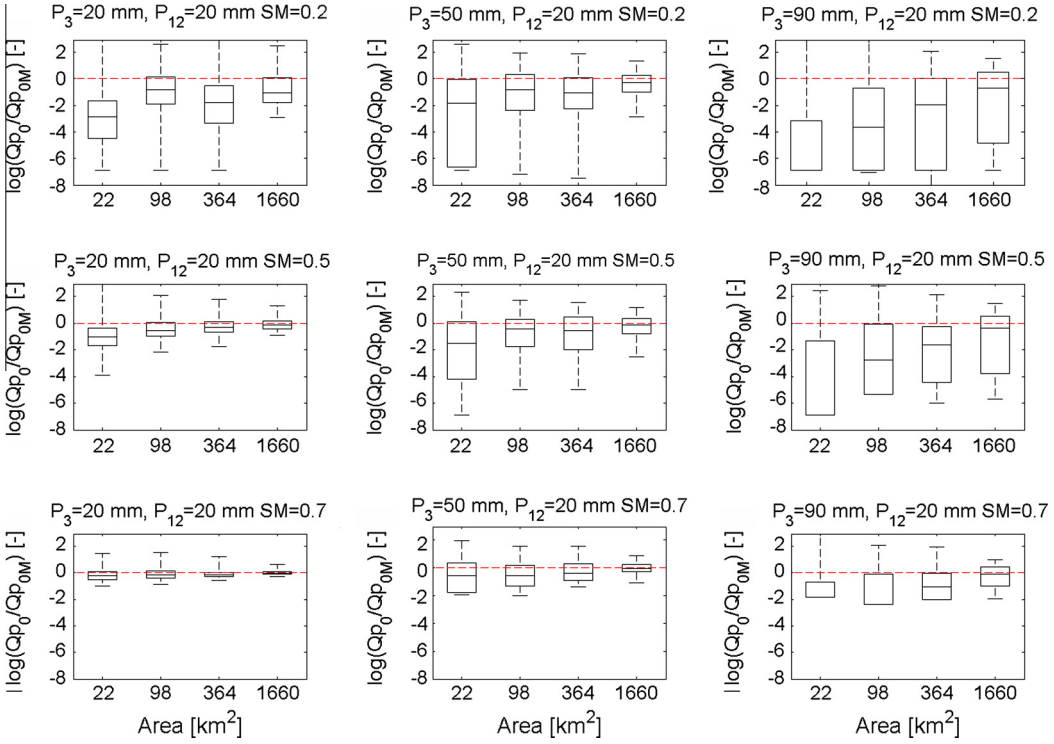

Fig. 10. Box plots of the peak flows (Qp_0) normalized with the average of the peak flows (Qp_{0M}) as a function of basin area for different values of P_3 and different values of SM. P_{12} = 20 mm. Log scale is used on *y*-axis.

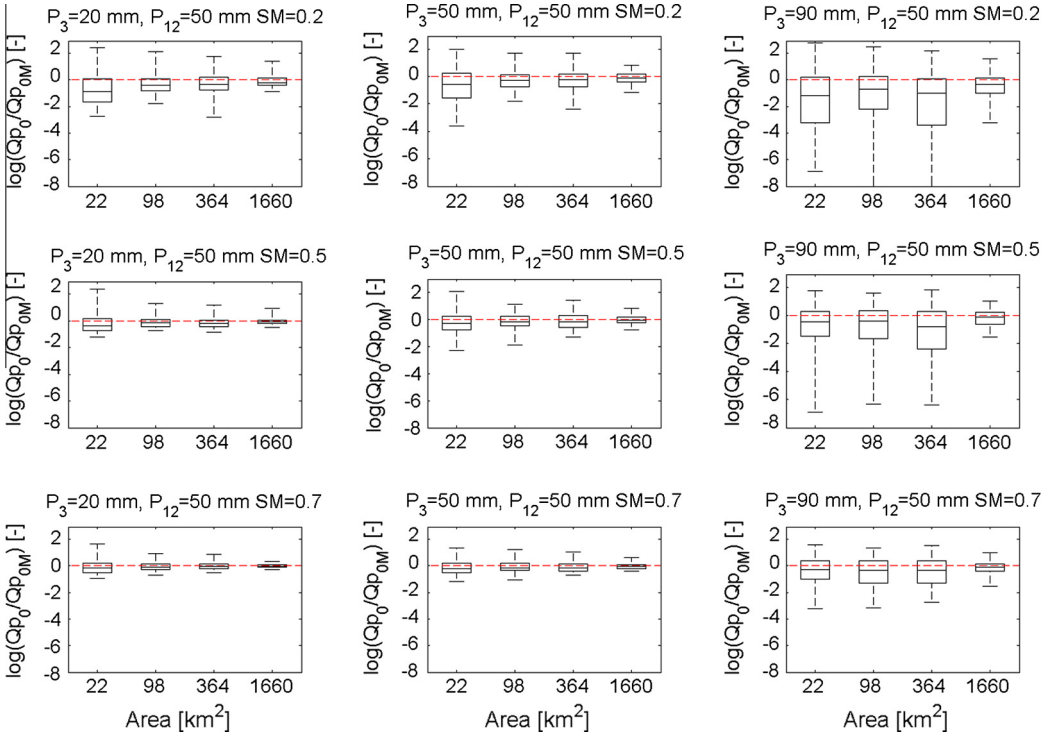

Fig. 11. Same as figure 10 with P_{12} = 50 mm.

change depending on the reliability of the precipitation forecast. To answer this question, three significant P_{12} values were considered and the probability curves were then built.

Fig. 7 reports the curves for the considered P_{12} errors and for the three values of SM. A sub-set of significant P_{12} values has been individuated. Err + 0 indicates the perfect forecast while the other symbols refer to the forecast with a certain error applied to P_{12}. On the *x*-axis the return period is reported, while the *y*-axis reports the probability that in at least one basin of the homogeneous region the return period *T* is exceeded. In the Liguria Region, a frequency analysis of the peak discharges based on a regional approach is available (Boni, 1999; Boni et al., 2007). This kind of approach

allows peak flows for fixed return periods (even for high return periods) to be estimated for each basin of the region based on basin characteristics (i.e. drainage area) and on a regional growth curve. In Table 2, the peak flows for fixed return periods are reported for the basins belonging to alert sub-region *C*.

As expected, for a fixed P_{12} and a fixed error on P_{12} (Err on Fig. 7), the probability for a fixed T increases with the increasing of SM. This is more evident in the cases where P_{12} = 50 mm and P_{12} = 100 mm.

The errors in P_{12} are reflected on the probability curves, and in many cases the multi-catchment approach leads to the amplification of the errors from precipitation to flood forecast; as an example for P_{12} = 150 mm, T = 50 yrs, SM = 0.5, we have $P(\%) = 0.5$ for the perfect forecast while $P(\%)$ about 0.2 and 0.9 for errors of ±30 % on P_{12} respectively. This is due to the way the curves are built. Notwithstanding, there is an interesting consideration. Consider, for example, the case of P_{12} = 100 mm and T = 10 yrs, it is evident that the curves are quite different depending on the applied error; this is especially true for errors of the order of 20–30%. However, even for a significant underestimation of P_{12}, the probability does not drop to 0. This is essential regarding the importance of using the multi-catchment approach when small-basins are considered: the peak flow with T = 10 yrs is not reached in many basins when single-site forecast is considered, but it can occur somewhere in the region.

To summarize the results we can state that: (i) the probability of exceedance of a fixed T increases with both the increasing of P_{12} values and SM values that means wetter initial condition, (ii) the analysis evidences an amplification of the errors on P_{12} for all the SM conditions; the multi-catchment procedure counteracts the decrease of variability with increasing values of SM evidenced in the case of single-site approach, (iii) the multi-catchment approach allows to moderate the underestimation of the flood forecast at sub-region scale even when relevant errors are applied to P_{12}.

4.2. Experiment 2

In this section, the impact of P_3 combined with P_{12} and SM is analyzed. Again, figures will be used to present the results.

For simplicity we considered three values of P_{12}: 20, 50, and 100 mm. The chosen P_{12} values are representative of low, medium and high amounts of precipitation and based on the experience of the authors and of the personnel of the FFC and on historical events.

4.2.1. Single site

Fig. 8 reports three subplots, one for each of the three P_{12} values considered. For each set of P_{12}, P_3 and SM, the flood forecast using N_e = 400 was done. The x-axis reports the value of P_3, while the y-axis reports the value of the corresponding mean peak flow. An increase of the P_3 corresponds to an increase of peak flow for the same P_{12}. In the case of P_{12} = 20 mm, there is a decrease for P_3 > 80 mm because the P_3 and P_{12} values are not consistent anymore (total volume of precipitation related to P_3 is greater than total volume related to P_{12}).

In the case of P_{12} = 100 mm, the increase is less evident, we should therefore consider P_3 values larger than 100 mm that are quite rare, even though physically possible in this region (Silvestro et al., 2012; Rebora et al., 2013).

An increase of SM causes a general increase of Q_p for fixed P_{12} and P_3 values. The graph shows evidence that the difference can be very large (100–300%) when changing the SM from 0.2 to 0.7, thus confirming the great impact of the initial soil moisture conditions on the flood forecast.

The graph also shows that even in the case of a good P_{12} forecast, a wrong evaluation of the type of event regarding the small spatial and temporal scales (i.e. P_3) can lead to a bad flood forecast, with underestimation (or overestimation) of the peak flows.

Fig. 9 reports the box plot of the peak flows of each ensemble normalized with the mean peak flow of the ensemble itself. This helps to evaluate the variability of the forecast in the different cases. In this case, the logarithmic scale is utilized for the y-axis.

Generally, the variability increases with an increase in the value of P_3 and decreases for wetter SM conditions. When the soil is dry (SM = 0.2) the trend is more complex (see subplots for SM = 0.2 in Fig. 9), with high variability for low values of P_3; the box initially decreases until P_3 reaches a certain value and then it starts to

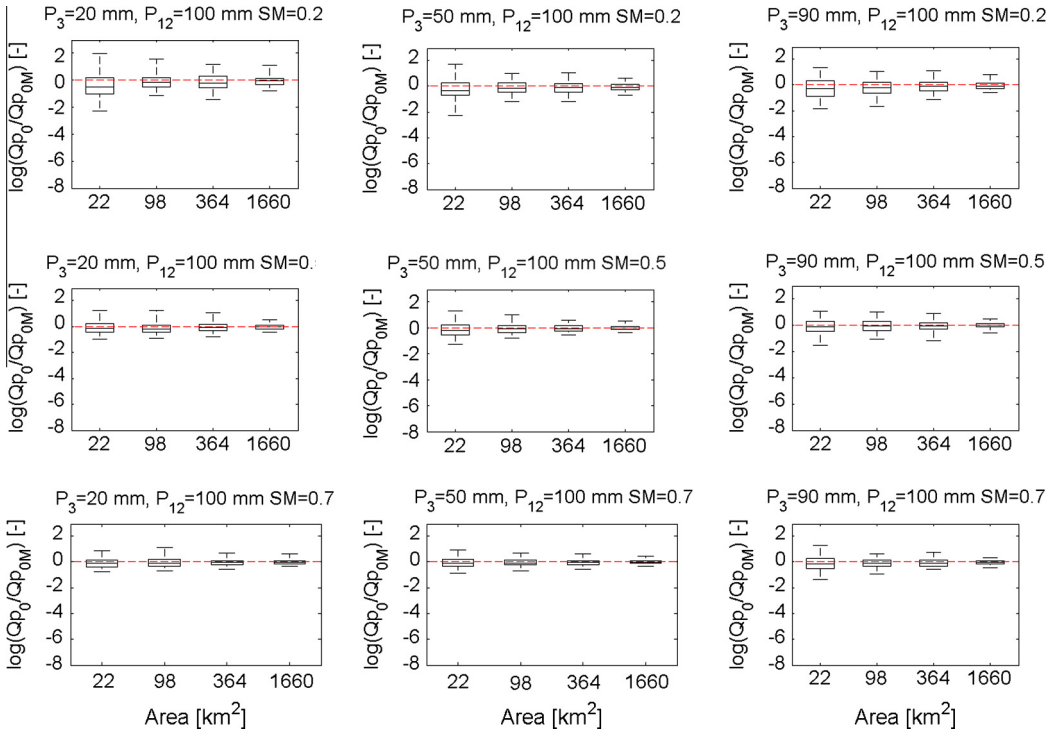

Fig. 12. Same as figure 10 with P_{12} = 100 mm.

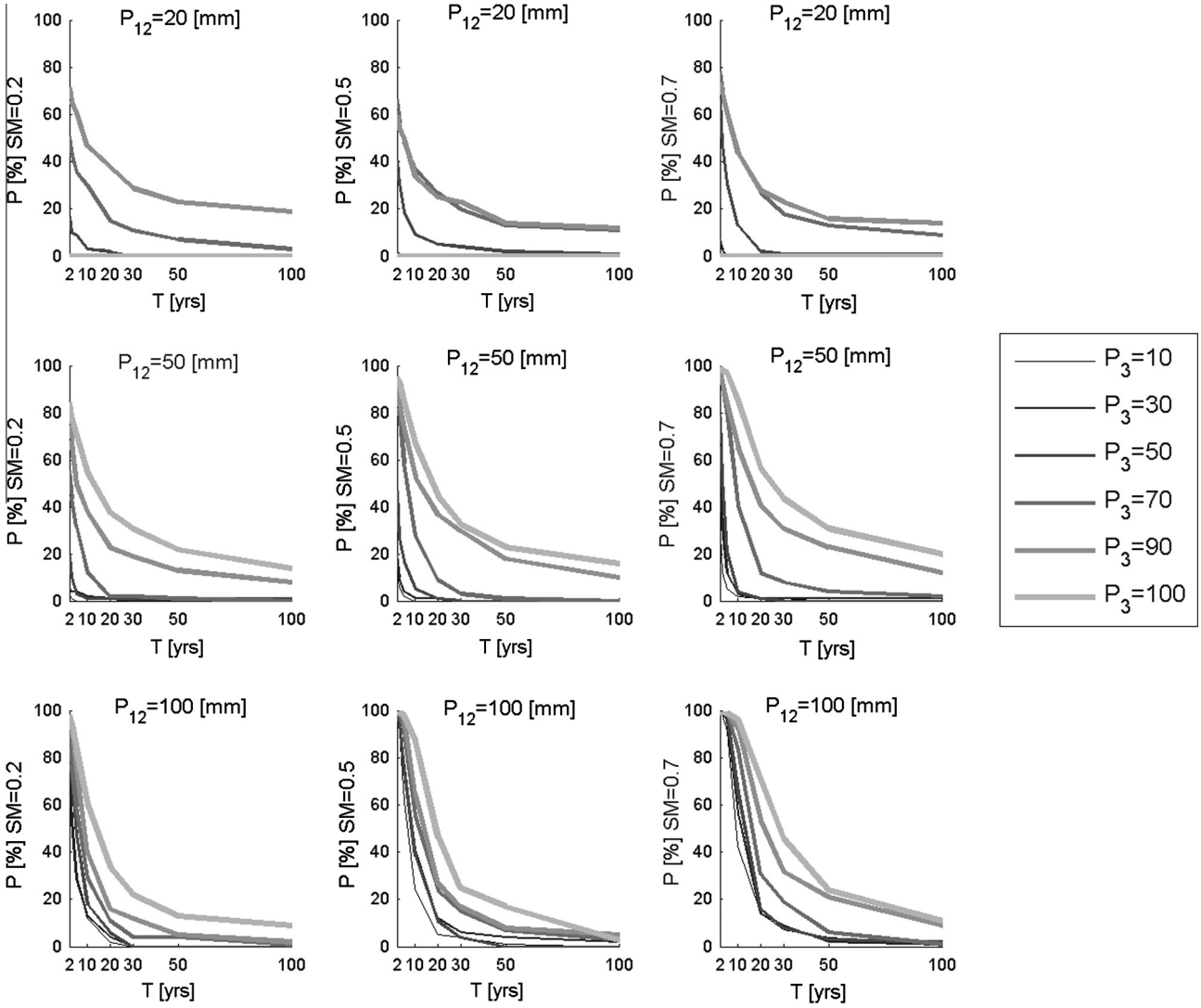

Fig. 13. Probability of exceedance curves for sub-region C as a function of P_{12} of P_3 and of the soil moisture conditions, SM.

increase. This is due to the strong impact of the runoff process with low values of the peak flows which has high variability depending on the spatial distribution of the rainfall scenarios (the same amount of rainfall on different types of soil could result in significantly different runoff). After certain values of P_3 are reached, the downscaling procedure generates very concentrated and intense events with possibly larger peak flows that lead again to an augmentation of the variability. In the case of $S = 0.7$, the variability increases for increasing values of P_3, this is more evident in the cases where $P_{12} = 50$ mm and $P_{12} = 100$ mm.

Similar to what was done in Experiment 1, we considered some significant values of P_{12} and P_3 in order to show what happens in basins with different drainage areas (Figs. 10–12). Again, the logarithmic scale is considered for the y-axis. We can summarize the behavior with the following three main points: (i) variability increases when the ratio between P_3 and P_{12} increases; (ii) variability decreases with the increasing of the basin area; (iii) variability decreases for wetter initial conditions, that is, when SM increases.

4.2.2. Multi-catchment

Interesting results are obtained by changing the P_3 values for the same value of P_{12}. The results are summarized in the multi-plot of Fig. 13. For fixed values of P_{12} and SM, the exceedance probability curves corresponding to different P_3 values are plotted on the same graph. From the graph, it is clear that the exceedance probability increases with the increasing of P_3, this again provides evidence that furnishing a good forecast of the total volume of precipitation on the sub-region (P_{12}) is not enough to obtain a reliable flood forecast. The characterization of the event type, so a good prediction of P_3 value, is crucial.

Even in this case, as for experiment 1, for fixed values of P_{12} and fixed P_3, the probability for a fixed T increases with the increasing of SM. This is more evident for the cases where $P_{12} = 50$ mm and $P_{12} = 100$ mm.

In the case where $P_{12} = 20$ mm, $P(T) = 0$ in some cases. This occurs for two reasons: (i) in some cases, the forecasted streamflow is low for all the ensemble members, this occurs for low P_3 values; (ii) in other cases, the condition of Eq. (1) is not satisfied ($P_3 = 100$ mm).

To summarize the results we can state that: (i) as in the case of experiment 1 the probability of exceedance of a fixed T increases for increasing values of P_{12} and SM, (ii) the probability of exceedance of a fixed T increases for increasing values of the ratio P_3/P_{12}, (iii) P_3 results to be a crucial element that can dramatically change the results even in the case of a P_{12} good prediction.

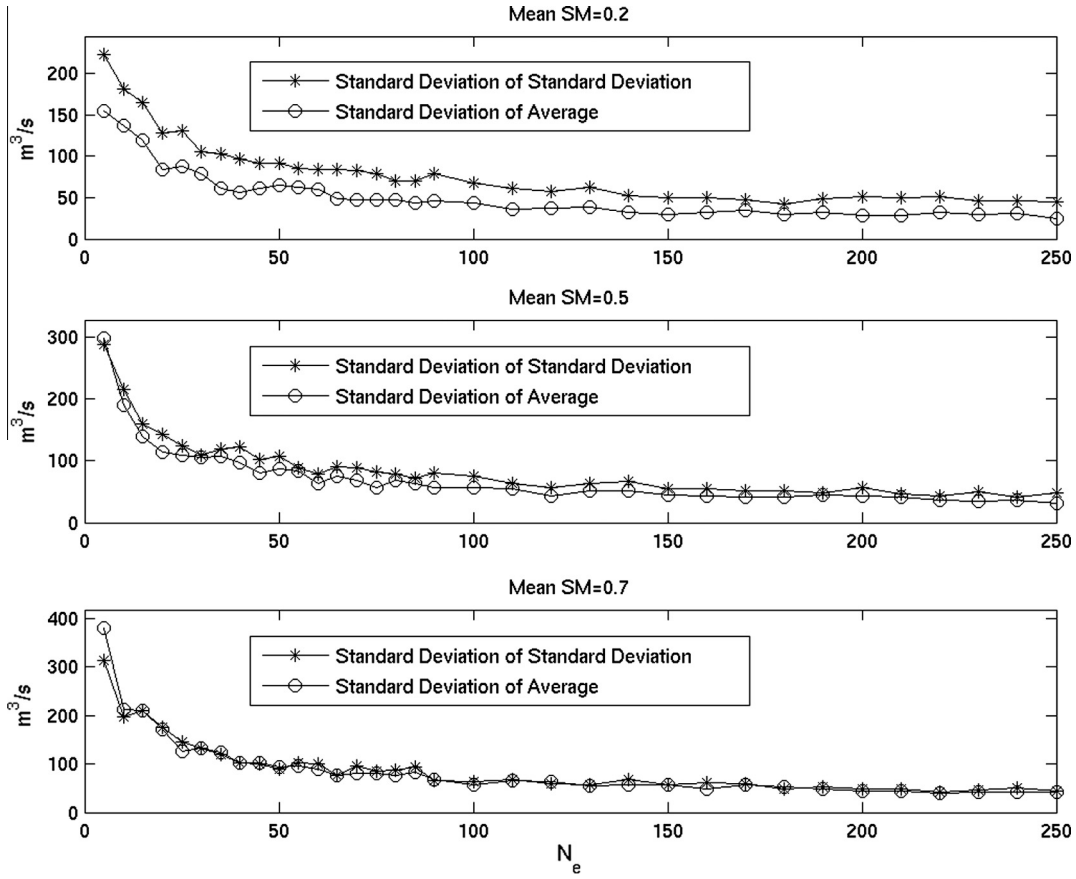

Fig. 14. Reference basin (364 km^2). Trend of σQM$_e$ and $\sigma(\sigma Q_e)$ as a function of the ensemble size, N_e.

4.3. Experiment 3

As introduced in Section 3.3, the P_{12} and P_3 values to carry out this experiment have been chosen based on the results of Experiment 2. In particular, analyzing Figs. 8 and 9, one can deduce that a P_{12} = 50 mm with P_3 in the range 70–90 mm, can produce sufficiently high peak flow values (Fig. 8) with a high variability (Fig 9); therefore, P_3 = 90 mm was chosen as the final value for this Experiment.

4.3.1. Single site

To show the impact of the ensemble size N_e on a single site forecast, the graphs in Fig. 14 were created. They have N_e on the x-axis while on the y-axis they have: (i) the standard deviation of the mean peak flows of the M forecast realizations for each ensemble size N_e (σQM$_e$); (ii) the sample standard deviation of the standard deviation of the peak flows of the M forecast realizations for each ensemble size N_e ($\sigma(\sigma Q_e)$).

These two statistics, as expected, decrease with the increase of N_e, in fact, we tend to simulate the rainfall scenarios of "all the possible worlds" with the downscaling procedure in each one of the M realizations of the forecast. As a consequence, when N_e assumes high values, the σQM$_e$ and $\sigma(\sigma Q_e)$ for sizes N_e, $N_e - 1$, $N_e + 1$,...etc. do not change significantly. This means that the results of the flood forecast are statistically the same. Theoretically, the best approach is using N_e as large as possible, but the three subplots show that for N_e values between 50 and 100 the derivative of the graphs drops to $\sim$0. This means that in an operational context the aforementioned range is sufficient to represent the variability of the forecast and to reduce to a negligible level the impact of the ensemble size on the final results. In summarizing the results, we could state that for $N_e > N_{et}$ (with $N_{et} \cong 50 \div 100$) the ensemble size negligibly affects the results.

Both σQM$_e$ and $\sigma(\sigma Q_e)$ increase with SM for fixed values of N_e, but no significant influence of the SM initial conditions on the N_{et} value seems to arise; it is, in fact, similar for all the considered SM values.

4.3.2. Multi-catchment

Interesting conclusions can be made regarding the impact of N_e on multi-catchment forecasts. It is not easy to visualize the results, but we attempted this with the help of several sub-plots (See Fig. 15). Each column represents a return period T, and has a subplot for every SM condition; each subplot has N_e on the x-axis and the box plot of probability of exceedance, P of the flow with return period T on the y-axis; in fact, each of the M realizations of the forecast of ensemble size N_e can produce a different P, it is therefore possible to build a box plot graph.

The figures are created for some representative return periods (T = 2, 20, 50 yrs). Looking at the case where T = 2 yrs, it is again clear that boxes (interquartiles) and whiskers have reduced amplitudes when N_e increases. The P for fixed T and fixed N_e increases when SM conditions become wetter. Consider the case where T = 2 yrs and SM = 0.7, there is always the certainty that T would be exceeded and therefore, the box-plot would not be necessary. The variability is larger for dry SM conditions and smaller for wet SM conditions.

This behavior changes when higher T values are considered, for example, when T = 20 yrs the variability increases from dry to wet conditions. For SM = 0.2, the considered pair of P_{12}, and P_3 causes the exceedance of T = 20 yrs for a reduced number of ensemble members; this happens for each of the M realizations. When SM

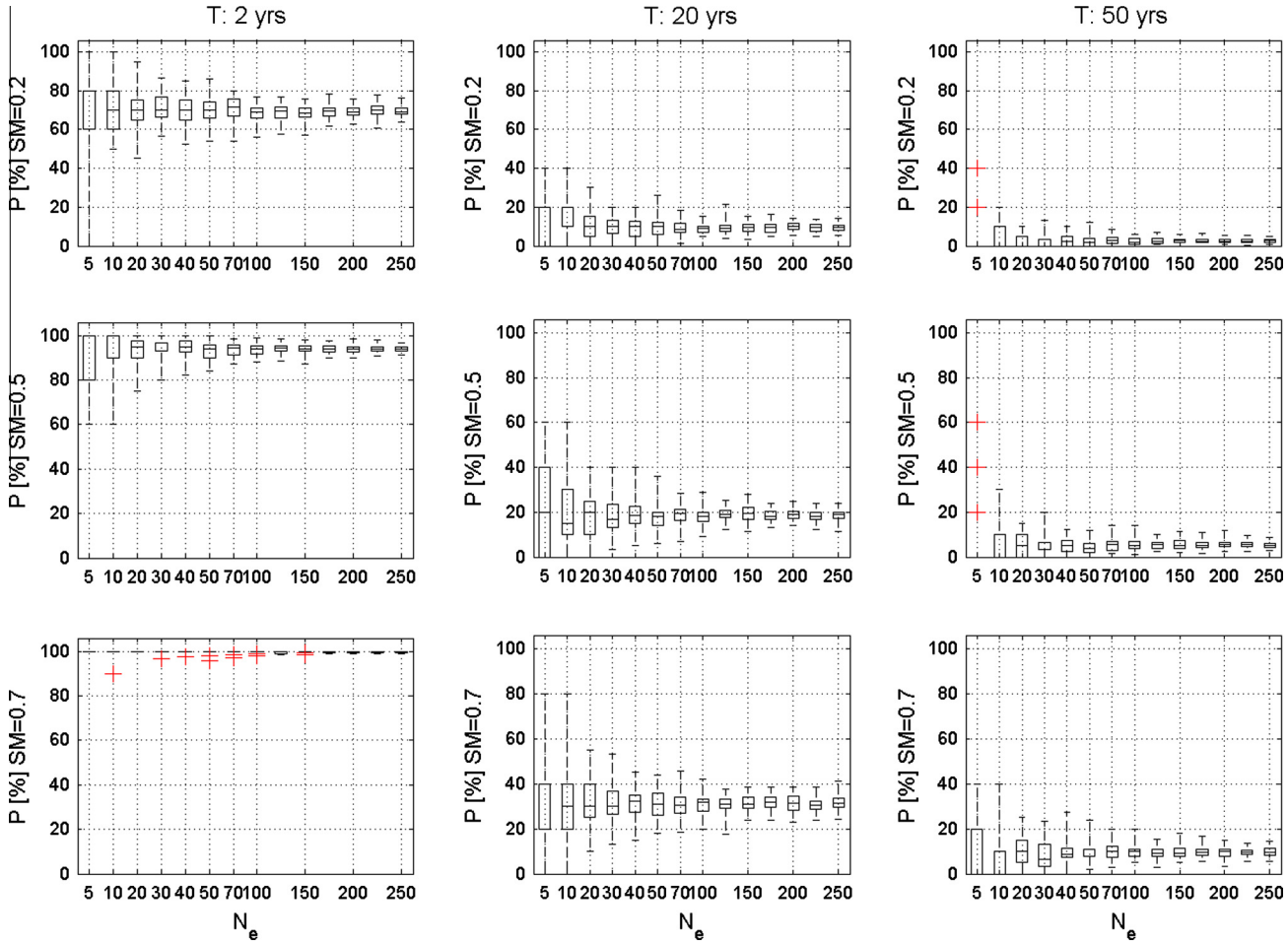

Fig. 15. T = 2, 20, 50 years; box plots of the probability of exceedance for sub-region C as a function of the ensemble size, N_e and the soil moisture condition, SM.

conditions are wetter, more rainfall scenarios can cause the exceedance of the considered T and the variability becomes higher. This is an interesting effect that is not totally intuitive, in fact, depending on the P_{12} and P_3 values, the uncertainty in the results of the multi-catchment approach increases from dry to wet SM conditions. Considering subplots for T = 20 yrs and T = 50 yrs, one can deduce that it would be necessary to increase the ensemble members N_e for SM = 0.7 in order to produce boxes that have the same dimensions as in the case of SM = 0.2 (i.e. the same degree of variability).

The main findings for the multi-catchment approach are the following: (i) the probability of exceedance P for fixed T and fixed N_e increases when SM increases, (ii) the variability of P between the M realizations of the flood forecast decreases when N_e increases, as a consequence the M realizations are statistically indistinguishable for large N_e, (iii) the behavior of the variability of P with SM is not always easy to be interpreted: depending on the considered T it can increase or decrease when SM increases.

In the case of the multi-catchment analysis, we suggest that the minimum ensemble size N_{et} be set to a value of at least 100.

5. Discussion and conclusions

The presented work aims at analyzing the propagations of the errors of an "Expert Quantitative Precipitation Forecast" on a probabilistic streamflow forecast. This "expert" forecast is mainly made-up of two quantities: a value that represents the total rainfall on predefined regions and temporal window (P_{12}) and a value that accounts for the rainfall intensities at small spatial and temporal scales (P_3). Precipitation uncertainties were combined with varying soil moisture initial conditions (SM).

The analysis was carried out considering the effects on a probabilistic forecast system operationally used at the Hydro-Meteorological Functional Centre of Liguria Region on both single-site (i.e. forecast on single basin) and multi-catchment (i.e. forecast on homogeneous regions) configurations. A wide range of EQPFs were synthetically generated and used as the input for a probabilistic flood forecasting system conceived for small and medium-sized basins made by a precipitation downscaling module and a semi-distributed hydrological model. Different initial soil moisture conditions were also considered as well as the impact of the ensemble size on the final flood forecast.

The experiments provided evidence that uncertainties related to the EQPFs combine with the non-linearity of the runoff process giving, in some cases, unexpected occurrences or results that are not immediately intuitive.

Errors on EQPFs are generally amplified in the case of dry SM conditions and "explosive" rainfall events with high convective characteristics tend to produce high variability in the flood forecasts. The non-linearity of the runoff process plays a fundamental role, especially when the rainfall events are severe, but not disastrous or very rare; in these cases, the rainfall amounts and intensities are large enough to generate critical streamflow values, but the latter are strongly influenced by the initial soil moisture conditions that can considerably change the final runoff amounts.

The results show that it is not simple to give a weight to the analyzed variables of the chain in order to individuate the major

cause of errors in the final flood forecasts. Even if the precipitation at the alert sub-region scale (P_{12}) plays a fundamental role because it fixes the total amount of precipitation volume, the soil moisture state and the type of the event synthesized by P_3 can influence in a substantial way the forecast. Looking at Fig. 8, it can be easily shown that, for example, for P_{12} = 50 mm and P_3 = 70 mm, the error on peak flow can be of the order of 200–400%, if SM are not correctly evaluated. On the other hand, for fixed P_{12} = 50 and SM, the peak flow can change by 2–4 times depending on the P_3 value. This, in such a way, formally confirms what was experienced in the operational use of the chain, that is, the fact that it is not easy to immediately identify the major sources of errors. Depending on the state of the soil and the type of event, one of the considered variables can assume a key role in order to perform a good or poor flood forecast.

In addition, the effects of the ensemble size, N_e on the flood forecast were considered. It is, in fact, not possible to consider all the possible "worlds" in operational applications of a probabilistic flood forecasting system, but N_e must be accurately set because as it was shown, it can influence the results. The experiments showed that after a certain value of N_e (that we named N_{et}), an increment in N_e has a negligible impact on the final results and does not introduce additional information about the variability of the forecasted streamflow values. The analysis showed that in the study area under consideration and for the considered forecast system, N_{et} assumes values ranging from 50 to 100.

Particularly interesting are the effects of the afore-mentioned combination of rainfall amount/intensity, SM initial condition and ensemble size on the results of the multi-catchment approach. The variability of the probability (P) of exceeding a certain return period, T on a considered region (for a fixed pair of values of P_3 and P_{12}) increases or decreases with SM conditions depending on what return period, T is considered.

The error model adopted in the presented analysis (uniformly distributed) is simple, this because the real distribution is unknown and it is due to very different uncertainty sources: NWPM errors, human subjective evaluation of the occurring event, different evaluation of the various NWPM reliability made by each forecaster. Because of these considerations the presented work could be even considered as a sensitivity analysis of the FFC in respect to P_3, P_{12} and SM. In particular the forecaster contribution to the prediction error is difficult to model because of its subjectivity. To make a rigorous analysis we should have the predictions of all possible events made by all forecasters available, but this is hard to be obtained and out of the scope of the work.

Of course it is possible that the choice of a different model error would lead to different results, for example differentiating the probability that the results due to a certain error occurs (in the study we assumed that an error of 5% or 30% on P_3 or P_{12}, as an example, has the same probability of occurrence). Anyway it is opinion of the authors that, at the state of the art, the assumption of more sophisticated error models is difficult to justify and these models would be not easy to parameterize.

The proposed analysis is tailored for probabilistic flood forecasting chains based on EQPF (Precipitation forecast on predefined areas and predefined time windows) and it is applicable on other systems that have such characteristic with slight modifications, anyway it could be adapted for being applied on a generic probabilistic/ensemble flood prediction system. The application of errors on precipitation forecast could be done even on rainfall fields derived by a NWPS, while the soil moisture initial condition is a parameter common to a wide number of hydrological models used in FFCs. Finally the analysis on the effects of the ensemble size can be carried out to a generic probabilistic/ensemble flood forecast system that uses a rainfall downscaling model.

As a final consideration, we conclude that forecasters must always exercise caution when interpreting the flood forecast system results; their variability is not constant, but depends on various factors such as the type and characteristics of the event and the soil moisture initial condition. The influence of the ensemble size on the results cannot be (totally) neglected, but setting it within a reasonable range can substantially reduce its impacts on the final forecast.

Acknowledgments

This work is supported by the Italian Civil Protection Department and by Liguria Region, Italy. We are very grateful to the meteorologists and the hydrologists of the Meteo-Hydrologic Centre of Liguria Region, for many useful discussions. We are grateful to Garvin Cummings for his suggestions in reviewing the quality of the writing.

References

Alfieri, L., Thielen, J., Pappenberger, F., 2012. Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland. J. Hydrol. 424–425, 143–153. http://dx.doi.org/10.1016/j.jhydrol.2011.12.038.

Arribas, A., Robertson, K.B., Mylne, K.R., 2005. Test of a poor man's ensemble prediction system for short-range probability forecasting. Mon. Wea. Rev. 133, 1825–1839.

Beven, K., Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Proc. 24, 43–69.

Boni, G., 2000. A Physically Based Regional Rainfall Frequency Analysis: Application to a Coastal Region in Northern Italy, In: EGS Plinius Conf. on Mediterranean Storms, Maratea, 14–16 October 1999, pp. 365–376.

Boni, G., Ferraris, L., Giannoni, F., Roth, G., Rudari, R., 2007. Flood probability analysis for un-gauged watersheds by means of a simple distributed hydrologic model. Adv. Water Resour. 30 (10), 2135–2144. http://dx.doi.org/10.1016/j.advwatres.2006.08.009.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., Bittelli, M., 2011. Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. Remote Sens. Environ. 115, 3390–3408.

Carpenter, T.M., Georgakakos, K.P., 2006. Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. J. Hydrol. 329, 174–185.

Ferraris, L., Rudari, R., Siccardi, F., 2002. The uncertainty in the prediction of flash floods in the northern Mediterranean environment. J. Hydrometeor. 3, 714–727.

Gabellani, S., Silvestro, F., Rudari, R., Boni, G., 2008. General calibration methodology for a combined Horton-SCS infiltration scheme in flash flood modeling. Nat. Hazards Earth Sci. 8, 1317–1327.

Germann, U., Berenguer, M., Sempere-Torres, D., Zappa, M., 2009. REAL: ensemble radar precipitation estimation for hydrology in a mountainous region. Q. J. Royal Meteorol. Soc. 135, 445–456.

Giannoni, F., Roth, G., Rudari, R., 2000. A semi – distributed rainfall – runoff model based on a geomorphologic approach. Phy. Chem. Earth 25 (7–8), 665–671.

Giannoni, F., Roth, G., Rudari, R., 2003. Can the behaviour of different basins be described by the same model_s parameter set? A geomorphologic framework. Phys. Chem. Earth 28, 289–295.

Liu, Z., Martina, M.L.V., Todini, E., 2005. Flood forecasting using a fully distributed model: application of the TOPKAPI model to Upper Xixian Catchment. Hydrol. Earth Syst. Sci. 9, 347–364.

Mascaro, G., Vivoni, E.R., Deidda, R., 2010. Implications of ensemble quantitative precipitation forecast errors on distributed streamflow forecasting. J. Hydrometeorol. 11, 69–86. http://dx.doi.org/10.1175/2009JHM1144.1.

Rebora, N., Ferraris, L., Hardenberg, J.H., Provenzale, A., 2006. The RainFARM: rainfall downscaling by a filtered auto regressive model. J. Hydrometeorol. 7 (4), 724–738.

Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, F., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., Parodi, A., 2013. Extreme rainfall in the Mediterranean: what can we learn from observations. J. Hydrometeorol. e-View. http://dx.doi.org/10.1175/JHM-D-12-083.1.

Siccardi, F., Boni, G., Ferraris, L., Rudari, R., 2005. A hydro-meteorological approach for probabilistic flood forecast. J. Geophys. Res 110, d05101. http://dx.doi.org/10.1029/2004jd005314.

Silvestro, F., Rebora, N., Ferraris, L., 2011. Quantitative flood forecasting on small and medium size basins: a probabilistic approach for operational purposes. J. Hydrometeorol. 12 (6), 1432–1446.

Silvestro, F., Gabellani, S., Giannoni, F., Parodi, A., Rebora, N., Rudari, R., Siccardi, F., 2012. A Hydrological Analysis of the 4th November 2011 event in Genoa. Nat. Hazards earth syst. Sci. 12, 2743–2752. http://dx.doi.org/10.5194/nhess-12-2743-2012.

Tramblay, Y., Bouaicha, R., Brocca, L., Dorigo, W., Bouvier, C., Camici, S., Servat, E., 2012. Estimation of antecedent wetness conditions for flood modelling in northern Morocco. Hydrol. Earth Syst. Sci. 16, 4375–4386.

Vrugt, J.A., Clark, P., Diks, C.G.H., Duan, Q., Robinson, D.A., 2006. Multi-objective calibration of forecast ensembles using Bayesian model averaging. Geophys. Res. Lett. 33, L19817. http://dx.doi.org/10.1029/2006GL027126.

Zappa, M., Jaun, S., Germann, U., Walser, A., Fundel, F., 2011. Superposition of three sources of uncertainties in operational flood forecasting chains. Atmos. Res. http://dx.doi.org/10.1016/j.atmosres.2010.12.005.